**Colorado Technical University**

 **Course:** MATH366 – Probability and Statistics

#### Unit 4 Part 07 Readings: Nonlinear Time Series and

#### Multiple Regression

**Regression**

 the technique of fitting a linear equation to data

 regression is used to: estimate or predict the value *y* (the response variable) for a

given *x* (the predictor variable)

 **equation of a regression line:**

y = b + mx

m is the slope of the equation

b is the y-intercept

y is the calculated y-value for a given x-value

 RSQ measures how accurately a regression line predicts the data

**Using regression equations**

 regression is used not only to fit data but to forecast (predict) observations -

especially for a time series

 for a time series, “time” is always the “x” (independent) variable

 the method of fitting the line is least squares (minimum vertical distance)

**Calculating values using a regression equation:** y = b + mx

 Suppose you had:

value = $15,000 - $500(age)

 To find the value if the age is 10, plug that value in to the “age” slot in the equation:

value = $15,000 - $500(10)

 = $15,000 -$5,000

 = $10,000

**Interpreting regression equations:**

 For the equation above:

 The value when the item is new (age 0) is b = $15,000

 The value decreases m = $500 each year the item ages

**Non-Linear Time series** – data that change over time

The goal is often to create an accurate forecast of future data values

Time series are frequently **non-linear**

Nonlinear time series can be nonlinear in a variety of ways:

One way to generate a non-linear regression is to square the

input (x) variable – this provides a parabolic curve fit



If you multiply the x2 variable by -1,

it will make a hill-shaped

fit line

A square root or log of x will provide a dampening increase

e-exponents provide an increasing curve



A cubic will fit a curve that levels off in the center



Sine curves provide an undulating curve fit

These are all called “curvilinear” trend lines

**Multiple Regression**

You can have as many independent variables as you want

A regression equation containing many independent variables ("x"s) is called a

"**multiple regression**"

Multiple regression equations are often called "**mathematical models**"

You can also have more than one dependent variable ("y"s)

A regression equation containing many dependent variables ("y"s) is called a

"**multivariate regression**"

In Excel, you can only have one “y”

Having more than one independent variable in a regression equation produces a

“**multiple R-squared**”

 The multiple R-squared is the RSQ in the regression -generated sheet